

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

CM nucleic - nucleic search, using sw model

Run on: August 17, 2004, 17:36:58 ; Search time 5616 Seconds

(without alignments)
11345.132 Million cell updates/sec

Title: US-09-725-010-1

Perfect score: 1470
Sequence: 1 ggaattcttcttctcattc.....cttttagtttcgtaataatg 1470

Scoring table: IDENTITY_NDC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

GenEmbl: 1: gb ba: 2: gb htg: 3: gb in: 4: gb om: 5: gb ov: 6: gb pac: 7: gb ph: 8: gb pl: 9: gb pr: 10: gb ro: 11: gb sts: 12: gb sy: 13: gb un: 14: gb vi: 15: em ba: 16: em fun: 17: em hum: 18: em in: 19: em mu: 20: em om: 21: em or: 22: em ov: 23: em pac: 24: em ph: 25: em pl: 26: em ro: 27: em sts: 28: em un: 29: em vi: 30: em htg hum: 31: em htg inv: 32: em htg other: 33: em htg mus: 34: em htg pin: 35: em htg rod: 36: em htg vrc: 37: em htg vrc: 38: em sy: 39: em htgo hum: 40: em htgo mus: 41: em htgo other:

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1470	100.0	1470	6	AX154641	AX154641 Sequence
2	505	34.4	2682	8	CA064206	U64206 Candida alb
3	153	10.4	2188	8	AF001978	AF001978 Candida a
4	136.2	9.3	42565	8	CAC35A5	AL033336 C.albican
5	91.2	6.2	86826	3	PFMAL3P5	AL034556 Plasmodiu
6	86.8	5.9	104892	2	AC005504	AC005504 Plasmodiu
7	86.8	5.9	169546	2	AC004157	AC004157 Plasmodiu
8	86.8	5.9	250421	3	AE014849	AE014849 Plasmodiu
9	77.6	5.3	105682	3	AC116977_3	Continuation (4 of
10	75.2	5.1	302156	3	AC116977	AC116977 Dictyoste
11	74.6	5.1	343050	3	PFPA929353	AF538053 Monosiga
12	74	5.0	76568	3	MBREV	AE014826 Plasmodiu
13	73.8	5.0	250663	3	AE014826	AE014826 Plasmodiu
14	73.4	5.0	14867	3	AE001398	AE001398 Plasmodiu
15	73.4	5.0	164399	3	PFMAL3P6	Z98551 Plasmodiu
16	72.8	5.0	80803	9	AC133522	AC133522 Homo sapi
17	72.8	5.0	113880	3	PFMAL3P4	AL008970 Plasmodiu
18	72.6	4.9	110000	2	PFMAL7P1_07	Continuation (8 of
19	72	4.9	8056	6	AX559046	AX559046 Sequence
20	72	4.9	349751	3	PFMAL4P3	AL035476 Plasmodiu
21	71.8	4.9	145892	9	AC134919	AC134919 Homo sapi
22	71.8	4.9	343050	3	PFPA929353	AL929353 Plasmodiu
23	71.4	4.9	25568	3	AC116955	AC116955 Dictyoste
24	71.4	4.9	39984	3	AC116101	AC116101 Dictyoste
25	70.8	4.8	2009	6	AX457067	AX457067 Sequence
26	70.8	4.8	271546	3	AE014843	AE014843 Plasmodiu
27	70.4	4.8	192929	2	AC005505	AC005505 Plasmodiu
28	70.4	4.8	250713	3	AE014850	AE014850 Plasmodiu
29	70.2	4.8	281723	3	PFPA929359	AL929359 Plasmodiu
30	69.4	4.7	182870	3	AC116960	AC116960 Dictyoste
31	69.2	4.7	253305	3	PFMAL3P7	AL034559 Plasmodiu
32	68.6	4.7	166719	9	AC136957	AC136957 Homo sapi
33	68.6	4.7	198434	9	AC109994	AC109994 Homo sapi
34	68.2	4.6	192187	3	AC116920	AC116920 Dictyoste
35	68	4.6	9810	6	AX345328	AX345328 Sequence
36	67.8	4.6	250029	3	AE014830	AE014830 Plasmodiu
37	67.6	4.6	170966	5	BX000991	BX000991 Zebrafish
38	67.4	4.6	66993	2	AC138074	AC138074 Homo sapi
39	67.2	4.6	1708	8	AF500786	AF500786 Plasmocyes
40	67.2	4.6	7347	1	AF211124	AF211124 Caronell
41	67.2	4.6	67970	3	PFMAL3P1_04	AL031746 Plasmodiu
42	67.2	4.6	110000	2	PFMAL3P1_04	Continuation (5 of
43	67	4.6	8056	6	AX559046	AX559046 Sequence
44	67	4.6	110000	2	AL954295_2	Continuation (3 of
45	67	4.6	122838	2	BX248494	BX248494 Danio rer

ALIGNMENTS

RESULT 1
AX154641
LOCUS AX154641
DEFINITION Sequence 1 from Patent WO0138550.
ACCESSION AX154641
VERSION AX154641.1 GI:14536200
KEYWORDS
SOURCE
ORGANISM
Candida albicans
Fukariyocai; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE
AUTHORS
TITLE
Methods for altering the expression of hyphal-specific genes

FEATURES
source

Sundstrom, Paula (US)
Location/Qualifiers
1. 1470
/organism="Candida albicans"
/mol_type="unassigned DNA"
/db_xref="taxon:5476"

ORIGIN

Query Match 100.0%; Score 1470; DS 6; Length 1470;
Best Local Similarity 100.0%; Pred. No. 3.6e-226;
Matches 1470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GATCTTTCTTTTCATTCCCTTAAACCGATCAGAGAAAGAGAGAAATTAAGCTAT 60
DB 1 GGATCTTTCTTTTCATTCCCTTAAACCGATCAGAGAAAGAGAGAAATTAAGCTAT 60
QY 61 GATTAATGTTGATTTGTATTAATCAATCACTAAGACGCTTGACAGTTAAAGTACG 120
DB 61 GATTAATGTTGATTTGTATTAATCAATCACTAAGACGCTTGACAGTTAAAGTACG 120
QY 121 TTGTTGTGTCCTGCTCGTCTAATTTCTGTGACAGAGATTATTAACAAGAAATACG 180
DB 121 TTGTTGTGTCCTGCTCGTCTAATTTCTGTGACAGAGATTATTAACAAGAAATACG 180
QY 181 GAAACCCCTCCAAAATAAATTTTGAACCTTACACGACATTAATTCGGATTAACCTTCG 240
DB 181 GAAACCCCTCCAAAATAAATTTTGAACCTTACACGACATTAATTCGGATTAACCTTCG 240
QY 241 CATATATAAACTCTTGAAACATACGATATGTTATTTCTTCAATCTGGAATATTTT 300
DB 241 CATATATAAACTCTTGAAACATACGATATGTTATTTCTTCAATCTGGAATATTTT 300
QY 301 GCTTTTCTTTTAACTATAGAACATTTGAAAAAAGGAAATGAAAAGTAAAGTGGCC 360
DB 301 GCTTTTCTTTTAACTATAGAACATTTGAAAAAAGGAAATGAAAAGTAAAGTGGCC 360
QY 361 TAACCATGAAAAATTAAGGCTAAGGTTTTCCTGATCGCTTAACTAAAAAGGAAATAA 420
DB 361 TAACCATGAAAAATTAAGGCTAAGGTTTTCCTGATCGCTTAACTAAAAAGGAAATAA 420
QY 421 CAAGAATTTATAGCATTAACCTGCTAAGGTCTCAACAAATATATTTTGAACGTTAGCT 480
DB 421 CAAGAATTTATAGCATTAACCTGCTAAGGTCTCAACAAATATATTTTGAACGTTAGCT 480
QY 481 CAAAGATTTATAGCATTAACCTGCTAAGGTCTCAACAAATATATTTTGAACGTTAGCT 540
DB 481 CAAAGATTTATAGCATTAACCTGCTAAGGTCTCAACAAATATATTTTGAACGTTAGCT 540
QY 541 CTCTCAGCTGACTGAATTTATCCATCTGAATTTATCACTCACTATATTCATTAATAA 600
DB 541 CTCTCAGCTGACTGAATTTATCCATCTGAATTTATCACTCACTATATTCATTAATAA 600
QY 601 TAGATTAAGTATTTGCTCTCTCACTAATTAATTAATTAATTAATTAATTAATTAAT 660
DB 601 TAGATTAAGTATTTGCTCTCTCACTAATTAATTAATTAATTAATTAATTAATTAAT 660
QY 661 TTGATTAAGTATTTGCTCTCTCACTAATTAATTAATTAATTAATTAATTAATTAAT 720
DB 661 TTGATTAAGTATTTGCTCTCTCACTAATTAATTAATTAATTAATTAATTAATTAAT 720
QY 721 GTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
DB 721 GTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
QY 781 AAACATTAACCTTTGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 840
DB 781 AAACATTAACCTTTGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 840
QY 841 CAAATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
DB 841 CAAATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
QY 901 TACTATTAAGTCAAACTAATTAATTTCTTCAAAAATGTTATTAATTAACAAGTCAATTAAT 960
DB 901 TACTATTAAGTCAAACTAATTAATTTCTTCAAAAATGTTATTAATTAACAAGTCAATTAAT 960

```

```

DB 901 TACTATTAAGTCAAACTAATTAATTTCTTCAAAAATGTTATTAATTAACAAGTCAATTAAT 960
QY 961 TCTTGGATTCGAAAACAGAAATTCGAAATTCGAGATTAATGTCAGTCACAATTC 1020
DB 961 TCTTGGATTCGAAAACAGAAATTCGAAATTCGAGATTAATGTCAGTCACAATTC 1020
QY 1021 ATTGTAAGGAGAGATTTGTGTGCTCATATCCGCTTATTAATGTAACGCTTAAGTAA 1080
DB 1021 ATTGTAAGGAGAGATTTGTGTGCTCATATCCGCTTATTAATGTAACGCTTAAGTAA 1080
QY 1081 TCTAATAACACACACCTTTCTAATAACCTAATTAATTAATTAATTAATTAATTAATTA 1140
DB 1081 TCTAATAACACACACCTTTCTAATAACCTAATTAATTAATTAATTAATTAATTAATTA 1140
QY 1141 GATTAAGTATGTTAGCCGCTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1200
DB 1141 GATTAAGTATGTTAGCCGCTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1200
QY 1201 CACTTTTGTGCGACTTATTAATACGCTTTTGAACACTCTCTTTGATACCTGATACCG 1260
DB 1201 CACTTTTGTGCGACTTATTAATACGCTTTTGAACACTCTCTTTGATACCTGATACCG 1260
QY 1261 CCTTTTCTTAACTAGCACTCTGTAAGAGTCCCTTTCTTTTCTTTTCTTTTCTTTTCTTT 1320
DB 1261 CCTTTTCTTAACTAGCACTCTGTAAGAGTCCCTTTCTTTTCTTTTCTTTTCTTTTCTTT 1320
QY 1321 TGAATATGTAATACAGATTTTCTTCAAAAATTAATTAATTAATTAATTAATTAATTAAT 1380
DB 1321 TGAATATGTAATACAGATTTTCTTCAAAAATTAATTAATTAATTAATTAATTAATTAAT 1380
QY 1381 ATTTCATTTCCATTTCAACTGTTTCTTCAACATTAATTAATTAATTAATTAATTAATTA 1440
DB 1381 ATTTCATTTCCATTTCAACTGTTTCTTCAACATTAATTAATTAATTAATTAATTAATTA 1440
QY 1441 TAGTCACTGCTTTTGTGTTGTTGCTCATATG 1470
DB 1441 TAGTCACTGCTTTTGTGTTGTTGCTCATATG 1470

```

RESULT 2

CAU64206

2682 bp DNA linear PLN 17-DEC-1998

Candida albicans hyphal wall protein 1 (HWP1) gene, complete cds.

U64206 GI:4028879

ORGANISM

Candida albicans

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

Saccharomycetales; mitosporic Saccharomycetales; Candida.

REFERENCE

Staab,J.F., Ferrer,C.A. and Sundstrom,P.

Developmental expression of a tandemly repeated, proline-and

glutamine-rich amino acid motif on hyphal surfaces on Candida

albicans

J Biol. Chem. 271 (11), 6298-6305 (1996)

MEDLINE

J Biol. Chem. 271 (11), 6298-6305 (1996)

PUBMED

2 (bases 1 to 2682)

Staab,J.F. and Sundstrom,P.

Genetic organization and sequence analysis of the hypha-specific

cell wall protein gene HWP1 of Candida albicans

Yeast 14 (7), 681-686 (1998)

JOURNAL

8639315

3 (bases 1 to 2682)

Staab,J.F. and Sundstrom,P.

Direct Submission

Submitted (18-JUL-1996) Medical Microbiology & Immunology, Ohio

State University, 333 West 10th Avenue, Columbus, OH 43210, USA

4 (bases 1 to 2682)

REFERENCE

Staab,J.F.


```

RESULT 4
CAC35A5/c 42565 bp DNA linear PLN 05-NOV-1998
LOCUS C.albicans cosmid Ca35A5.
DEFINITION Al033396
ACCESSION Al033396.1 GI:3850143
VERSION ARS; Canikl; CDP-alcohol phosphatidyltransferase; chitin synthase;
KEYWORDS chsi; cytochrome P450; DNA polymerase delta; histidine kinase;
isocitrate dehydrogenase pseudogene; LTR; pol3; Rcc1; regulator of
chromosome condensation; rehydrin; Rps1-like region.
SOURCE Candida albicans
ORGANISM Candida albicans
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE 1 (bases 1 to 42565)
AUTHORS Tait, B., Simon, M.C., King, S., Brown, A.J., Gow, N.A. and Shaw, D.J.
TITLE A Candida albicans genome project: cosmid contigs, physical
mapping, and gene isolation
JOURNAL Fungal Genet. Biol. 21 (3), 308-314 (1997)
MEDLINE 97435544
PubMed 9290243
REMARK Article No. FG970983
REFERENCE 2 (bases 1 to 42565)
AUTHORS Oliver, K. and Harris, D.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 42565)
AUTHORS Barrell, B.G. and Randle, M.A.
JOURNAL Direct Submission
Submitted (05-NOV-1998) On behalf of the pilot sequencing project
on the Candida albicans strain 1161 genome. Sanger Centre, Wellcome
Trust Genome Campus, Hinxton, Cambridgeshire CB10 1SA E-mail:
barrell@sanger.ac.uk cosmids supplied by Prof. Duncan Shaw [3]
Department of Molecular and Cell Biology, The Institute of Medical
Science, University of Aberdeen, Foresterhill, Aberdeen, AB25 2ZD,
UK
COMMENT
Notes:
Funding: sequencing funded by Beowulf Genomics Ltd. CDS are
numbered using the following system eg CAC20C1.01c. CA (C.
albicans), 20C1 (cosmid name), 01 (complementary strand).
The more significant matches with motifs in the PROSITE/PRM
database are also included but some of these may be fortuitous.
IMPORTANT: This sequence MAY NOT be the entire insert of the
sequenced clone. It may be shorter because we only sequence
overlapping sections once, or longer, because we arrange for a
small overlap between neighbouring submissions. Cosmid Ca35A5 is
likely to map to region R, chromosome 7.
LOCATION/Qualifiers
1..42565
/organism="Candida albicans"
/mol_type="genomic DNA"
/strain="1161"
/db_xref="taxon:5476"
/map="chromosome 7"
/clone="cosmid Ca35A5"
/complement(1..785)
/gene="Ca35A5.01c"
/complement(1..785)
/gene="Ca35A5.01c"
/more="Ca35A5.01c, partial orf, len: > 261 aa, most
similar to EPT1_YAST ethanolaminephosphotransferase (EC
2.7.8.1) (391 aa), fasta scores opt: 984, E(1): 0, (56.3%
identity in 263 aa overlap), contains P500379 CDP-alcohol
phosphatidyltransferases signature"
/codon_start=1
/transl_table=12
/product="putative alcohol phosphatidyl transferase"
/protein_id="CAA21944.1"
/db_xref="GI:3850144"
/db_xref="GOA:O94010"
/db_xref="SPTREMBL:O94010"
/translation="MGILFIPNKLQNLKYRSESDHSIISKYILKKNWPFVQIFPL

```

Fri Aug 20 10:28:07 2004

us-09-725-010-1.rng

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM nucleic - nucleic search, using sw model

Run on: August 17, 2004, 17:33:53 ; Search time 593 seconds

(without alignments)
10530.946 Million cell updates/sec

Title: US-09-725-010-1

Perfect score: 1470
Sequence: 1 ggaatctctcttcattc.....ctttagcttcgcatatcg 1470

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1459	99.3	1471	5	AAF90471	Aaf90471 Candida a
2	159.4	10.8	2093	3	AAAS7996	Aaas7996 2093 bp C
3	72	4.9	8056	7	ABZ10246	Abz10246 Haematopo
4	68	4.6	9810	6	ABZ10246	Abz10246 Human imm
5	67	4.6	8056	7	ABZ10246	Abz10246 Haematopo
6	66.4	4.5	18624	6	ABZ13157	Abz13157 Human imm
7	66	4.5	17538	6	ABZ13157	Abz13157 Human imm
8	65.6	4.5	7584	6	AAAS9251	Aaas9251 Long term
9	64.8	4.4	5884	6	ABZ13157	Abz13157 Human imm
10	64.4	4.4	34688	6	ABO67060	Abog67060 Human ang
11	63.4	4.3	5254	6	ABK31514	Abk31514 Signal tr
12	63.4	4.3	115218	7	ACA64845	Aca64845 Human HNR
13	62.8	4.3	15121	6	ABN80239	Abn80239 Human che
14	62	4.2	8056	7	ABZ10100	Abz10100 Haematopo
15	61.2	4.2	6775	6	ABO67159	Abog67159 Human ang
16	61.2	4.2	12592	6	AAAS61102	Aaas61102 Human gen
17	61	4.1	6620	4	AAAS4487	Aaas4487 Chemi call
18	61	4.1	6620	6	ABK28415	Abk28415 DNA trans
19	61	4.1	6620	6	ABN80278	Abn80278 Human che
20	61	4.1	8056	7	ABZ10100	Abz10100 Haematopo
21	60.2	4.1	12426	4	AAAS46482	Aaas46482 Tumour su
22	60	4.1	1501	7	ABZ10188	Abz10188 Haematopo
23	60	4.1	1501	9	ADBS4162	Adbs4162 Human lym

24	59.8	4.1	8310	2	AAZ29911	Aaz29911 cDNA enco
25	59.6	4.1	6494	6	ABU33392	Abu33392 Human imm
26	59.6	4.1	6494	6	ADP28390	Adp28390 Human che
27	59.2	4.0	3684	7	ACF62820	Acf62820 Colon can
28	59.2	4.0	3684	7	ABZ10128	Abz10128 Haematopo
29	59.2	4.0	6664	9	ADBS4322	Adbs4322 Pretreate
30	59.2	4.0	11334	6	ABU33233	Abu33233 Human imm
31	59.2	4.0	18154	6	ABU32255	Abu32255 Human imm
32	58.8	4.0	10048	6	ABU70314	Abu70314 Human gen
33	58.8	4.0	10048	6	AAAS61252	Aaas61252 Human gen
34	58.6	4.0	15373	6	ABU32467	Abu32467 Human imm
35	58.4	4.0	17131	6	ABU33053	Abu33053 Human imm
36	58.2	4.0	11049	6	ABU32669	Abu32669 Human imm
37	58.2	4.0	11049	6	ABU92219	Abu92219 Chemical1
38	58.2	4.0	11049	6	ABU49322	Abu49322 Human pol
39	58	3.9	5474	6	ABU33041	Abu33041 Human imm
40	58	3.9	5474	6	ABO66992	Abog66992 Human ang
41	58	3.9	7631	6	ABU32860	Abu32860 Human imm
42	58	3.9	13574	6	ABU33317	Abu33317 Human imm
43	57.8	3.9	5085	7	ACF62814	Acf62814 Colon can
44	57.8	3.9	8085	9	ADBS4280	Adbs4280 Pretreate
45	57.4	3.9	883	4	AA115210	Aa115210 Human bre

ALIGNMENTS

RESULT 1	AAF90471	standard; DNA; 1471 BP.
ID	AAF90471	
XX	AAF90471;	
AC	AAF90471;	
DT	06-AUG-2001	(first entry)
XX		
DE	Candida albicans HWP1 gene 5' upstream region.	
XX		
KM	Hypthal wall protein, HWP1; promoter; hypthal-specific gene; infection	
KX	candidiasis; therapy; antifungal; fungicide; ds.	
XX		
OS	Candida albicans.	
XX		
FH	Key	Location/Qualifiers
FT	promoter	1..1468
FT		/*tag= u
FT	protein_bind	complement(59..64)
FT		/*tag= a
FT	protein_bind	/bound_moiety= "NIT2"
FT		94..105
FT		/*tag= b
FT	protein_bind	/bound_moiety= "PHO4"
FT		complement(128..233)
FT		/*tag= c
FT	protein_bind	/bound_moiety= "NIT2"
FT		complement(265..270)
FT		/*tag= d
FT	protein_bind	/bound_moiety= "NIT2"
FT		complement(432..438)
FT		/*tag= e
FT	protein_bind	/bound_moiety= "NIT2"
FT		468..479
FT		/*tag= f
FT	protein_bind	/bound_moiety= "PHO4"
FT		470..478
FT		/*tag= g
FT	misc_feature	/function= "S.cerevisiae cAMP response element"
FT		560..565
FT		/*tag= h
FT	protein_bind	/bound_moiety= "NIT2"
FT		573..578
FT		/*tag= i
FT	protein_bind	/bound_moiety= "NIT2"
FT		705..710

```

/*tag= j
/bound_moiety= "NIT2"
complement (800. .805)
/tag= k
/bound_moiety= "NIT2"
807. .812
/tag= l
/bound_moiety= "NIT2"
921. .926
/tag= m
/bound_moiety= "NIT2"
1139. .1144
/tag= n
/bound_moiety= "NIT2"
1146
/tag= o
/ncore= "this base is not present in the Hmp1 gene 5'
flanking region sequence given in Fig 4 of the
specification"
1247. .1252
/tag= p
/bound_moiety= "NIT2"
1256. .1261
/tag= q
/bound_moiety= "NIT2"
1313. .1318
/tag= r
/bound_moiety= "NIT2"
1355. .1361
/tag= t
1422. .1427
/tag= s
/bound_moiety= "NIT2"

WO200138550-A2.

31-MAY-2001.

29-NOV-2000; 2000WO-US032464.

29-NOV-1999; 99US-0167672P.

(SUND/) SUNDSTROM P.

Sundstrom P;

WPI; 2001-367698/38.

interfering with expression of hyphal-specific genes in fungus for
inhibiting fungal cell growth involves interfering with transcription of
hyphal-specific genes mediated by cis acting sequences.

claim 21; Page 78-89; 95pp; English.
```

CC diabetic ketoacidosis, and patients in which the normal microbial flora
CC has been disrupted because of disease, trauma or chemical, radiation or
CC other immunosuppressive prophylaxis. Also provided is a method for
CC characterizing genes under control of a DNA binding protein

XX
XX Sequence 1471 BP, 509 A; 260 C; 186 G; 516 T; 0 U; 0 Other;

Query Match	99.3%;	Score 1459;	DB 5;	Length 1471;
Best Local Similarity	99.9%;	Pred. No. 3.1e-258;		
Matches 1470;	Conservative	0;	Mismatches 0;	Indels 1;
				Gaps 1

Qy	GGATCTTCCTCTTCACTTCCCTTAAACCGATCAAGAAAGAAAGTGAATTAAGCAT	60
Db	1 GGATCTTCCTCTTCACTTCCCTTAAACCGATCAAGAAAGAAAGTGAATTAAGCAT	60
Qy	GATTAATGTTGATTTGTGTATTCATCAATCAACTAGACGCTTGTACGCTTAAAAAGTAGC	120
Db	61 GATTAATGTTGATTTGTGTATTCATCAATCAACTAGACGCTTGTACGCTTAAAAAGTAGC	120
Qy	TTGTGTGTGTCTCGCTCGTCTATTTCTGTGACGAGGTAAATAACAAGAAATACAG	180
Db	121 TTGTGTGTGTCTCGCTCGTCTATTTCTGTGACGAGGTAAATAACAAGAAATACAG	180
Qy	181 GAAACCTCCAAAAAAAATTTTGGACCTTACACGACATTAATTGCGGATTAACCTTGC	240
Db	181 GAAACCTCCAAAAAAAATTTTGGACCTTACACGACATTAATTGCGGATTAACCTTGC	240
Qy	241 CAATAAATAAACCTCTTGAACATACACGATGTATCTTTCTCACTAATCGAAATATTTT	300
Db	241 CAATAAATAAACCTCTTGAACATACACGATGTATCTTTCTCACTAATCGAAATATTTT	300
Qy	301 GCTTTTCTTTTAAACATTAAGAACATTTGAAAAAAGAAAAAGTAAGGTAGCTTCC	360
Db	301 GCTTTTCTTTTAAACATTAAGAACATTTGAAAAAAGAAAAAGTAAGGTAGCTTCC	360
Qy	361 TAACTTGAATAATTAATAGCTTAAGTTTTCGTAGTCCGTTTAACTTAAAGGAAATA	420
Db	361 TAACTTGAATAATTAATAGCTTAAGTTTTCGTAGTCCGTTTAACTTAAAGGAAATA	420
Qy	421 CAAAGTTTATGCGATTAACCTGCTAAGTGTCAACAAATATATTTTCAAGTTAGCT	480
Db	421 CAAAGTTTATGCGATTAACCTGCTAAGTGTCAACAAATATATTTTCAAGTTAGCT	480
Qy	481 CTATGAAAAATTAACAACTAAATCTTAAAGAAATTCCTCTATATATATTAAGAAATCC	540
Db	481 CTATGAAAAATTAACAACTAAATCTTAAAGAAATTCCTCTATATATATTAAGAAATCC	540
Qy	541 CTCTCAGCTGAACTGAATTAATCCATCGAATTAACGTCCACATTAATCATCAATAAA	600
Db	541 CTCTCAGCTGAACTGAATTAATCCATCGAATTAACGTCCACATTAATCATCAATAAA	600
Qy	601 TAGATTAGTATGTTCTCTTCACTAACATTAACATTAATGCAATGCTAGCTATTC	660
Db	601 TAGATTAGTATGTTCTCTTCACTAACATTAACATTAATGCAATGCTAGCTATTC	660
Qy	661 TTGATTAATAGCATGTGACACACCTTAATTGAAACATTAACATTAATTAATTTCTT	720
Db	661 TTGATTAATAGCATGTGACACACCTTAATTGAAACATTAACATTAATTAATTTCTT	720
Qy	721 GTCTTCTTTTGTTTTCTTAACAAAAATGTTCCAGATTTTAAAAAATTTTGAAA	780
Db	721 GTCTTCTTTTGTTTTCTTAACAAAAATGTTCCAGATTTTAAAAAATTTTGAAA	780
Qy	781 AACAATAACCTTGAAGTATATATCAACTATTAACCTGTGTTTGAAGTAAAGAAAT	840
Db	781 AACAATAACCTTGAAGTATATATCAACTATTAACCTGTGTTTGAAGTAAAGAAAT	840
Qy	841 CAAATTTTCTTAACTGCACTAATGCACTTAACTGAACCTGATGTAATTTGCTATAC	900
Db	841 CAAATTTTCTTAACTGCACTAATGCACTTAACTGAACCTGATGTAATTTGCTATAC	900
Qy	901 TACTATTAAGTAAACAAATTAATCTTGAAGAAATGTTATATTAATCAAGCATATTAAT	960

Db	901	TACTATTAAGCCCAACAAATATATCTTTCAAATAAGTATATTAATTAACAAGTATCATATAT	960
Qy	961	TCTTTGGATCCAAAACAAAGAAATTCGGAAATTCAGACGATTAATATGCGATCCAAATTC	1020
Db	961	TCCTTTGGATCCAAAACAAAGAAATTCGGAAATTCGAGATTAATATGCGATCCAAATTC	1020
Qy	1021	ATTGTAAAAAGGAGAGATTGTTGGTAGGCTCAATATCGCTTAATATGTAACCTCTAAAGTAA	1080
Db	1021	ATTGTAAAAAGGAGAGATTGTTGGTAGGCTCAATATCGCTTAATATGTAACCTCTAAAGTAA	1080
Qy	1081	TCTTAAACAAACACAACTCTTAATAACCTTAATTAATTAACCTTAATGAGCTTCACACCGG	1140
Db	1081	TCTTAAACAAACACAACTCTTAATAACCTTAATTAATTAATTAACCTTAATGAGCTTCACACCGG	1140
Qy	1141	GATTA- GTTAGTTAGCCAGCGCTGTTTTTTTTTTGGCTTAATTTTAATGACTACATTTGTT	1199
Db	1141	GATTAATGTTAGTTAGCCAGCGCTGTTTTTTTTTTGGCTTAATTTTAATGACTACATTTGTT	1200
Qy	1200	TCACCTTTGTTGGCACTTTAAACCGTTTGGAACTTCCTGTGATCAGCTGATCC	1258
Db	1201	TGACTTTTGTGTGCACTTTAAATACCGTTTGTGCACTTCCTTTGTATGATCAGCTGATCC	1260
Qy	1260	GCGTTTTTTTAAACATGCAACTCTGTGTAAAGTCCCTTCTTTTCCCACTAATTTATCATTC	1319
Db	1261	GCGTTTTTTTAAACATGCAACTCTGTGTAAAGTCCCTTCTTTTCCCACTAATTTATCATTC	1320
Qy	1320	TTGAAAATATATATACAAATAGTTTTTTCAAAACATTAATTAATTAATTAATTAACCGG	1379
Db	1321	TTGAAAATATATATACAAATAGTTTTTTCAAAACATTAATTAATTAATTAATTAACCGG	1380
Qy	1380	TATTTTCAATTTCCATTCGAACTGTGTTTTCTCAACATATCAACACAAACAGGAATCTCT	1439
Db	1381	TATTTTCAATTTCCATTCGAACTGTGTTTTCTCAACATATCAACACAAACAGGAATCTCT	1440
Qy	1440	ATATGCACTCGCTTTTAGTTTCGTCAATATG	1470
Db	1441	ATATGCACTCGCTTTTAGTTTCGTCAATATG	1471

RESULT 2	
AAAS7996/c	
ID	AAAS7996 standard; DNA; 2093 bp.
AC	
XX	AAAS7996;
XX	
DT	10-OCT-2000 (first entry)
DB	2093 bp Candida albicans retrotransposon 15 sequence.
XX	
XX	Retrotransposon; pCal; Tca2; Ty1; copia; long terminal repeat; LTR;
KM	gas gene; group antigen; glycoprotein; pol; aspartate protease; integ-
KM	reverse transcriptase; RNaseH; pseudonoc; readthrough translation;
KM	stop codon suppression; gene delivery; gene therapy vector;
XX	genetic vaccine composition; immunogenic; transgenic animal; ds.
XX	
OS	Candida albicans.
PM	WO200026397-A1.
PD	
PD	11-MAY-2000.
PF	01-NOV-1999; 59WC-N2000179.
XX	
XX	30-OCT-1998; 98CA-02249046.
PR	30-OCT-1998; 98US-0106342P.
XX	
PA	(JANNC) JANSEN PHARM NV.
PI	Luyten WHUL, De Backer MD, Nelissen BJM, Poulter RTM;
XX	
DR	WPI; 2000-365640/31.
PT	Novel retrotransposon expression vectors useful for expressing an

PT antigen epitope or therapeutic agent, or detecting genes or the presence
 of *Candida* in a sample.
 XX
 XX Disclosure; Fig 71; 204pp; English.
 XX
 CC The invention relates to novel retrotransposons from the yeast *Candida*
 CC albicans which have a copy number of 40-150, preferably 50-100 copies per
 CC genome. In particular, the invention relates to the novel *C. albicans*
 CC Ty1/copia retrotransposon pCal (AA579320), and to the integrated form of
 CC this retrotransposon, designated TRG2, and to the novel *C. albicans*
 CC retrotransposons 1-28. pCal was initially isolated from *C. albicans*
 CC H031042 and has a copy number of 50-100 copies per cell. It comprises
 CC identical 280 bp long terminal repeats (LTRs) and two open reading frames
 CC (ORFs). The first ORF encodes a gag (group antigen) protein, and the
 CC second ORF encodes a polypeptide (pol) consisting of an aspartate
 CC protease, integrase, reverse transcriptase (RT) and RNaseH. The gag and
 CC pol ORFs of pCal are in the same reading frame, separated only by a
 CC termination codon (TGA). Translation of the pol ORF occurs through the
 CC occasional readthrough suppression of the stop codon, which is mediated
 CC by the formation of a pseudoknot within the gag-pol mRNA. The
 CC retrotransposons of the invention can be used as vectors for *in vitro* or
 CC *in vivo* transformation and expression. They can thus be used for the
 CC delivery and expression of a therapeutic, immunological or immunogenic
 CC molecule (e.g., an antigen) and may also be used for eliciting an
 CC immunological response in a host organism. They are therefore useful in
 CC genetic vaccine compositions and for gene therapy, particularly where the
 CC use of retroviral vectors is unsafe or undesirable. Additionally, the
 CC retrotransposons may be used to generate transgenic animals, to detect
 CC the presence of *Candida* in a sample, to detect and disrupt genes, and to
 CC assign functions to nucleotide sequences. Sequences AA57933-57936 and
 CC AA57982-58019 represent novel *C. albicans* retrotransposon sequences
 CC
 CC
 CC Sequence 2093 BP; 747 A; 332 C; 302 G; 712 T; 0 U; 0 Other;

	Query Match	Best Local Similarity	Score	159.4	DB 3	Length	2093
	Matches	191	Conservative	0	Mismatches	31	Indels
							Gaps
							1
Qy	480	TCTATAGAAATATATACAACTAAATCCCTTAAAGAAATTCCTCTATATATATATAGAAATC	539				
Db	1237	TCTATAGAAATATATACAACTAAATCCCTTAAAGAAATTCCTATATATATATAGAAATC	1178				
Qy	540	CCTCTCACAGTGAACACTGAATTCATCTGTAATTAACGTCGACTAATTC--CATCAAT	596				
Db	1177	CCTCTCACAGTGAACACTGAATTCATCTGTAATTAACGTCGACTAATTC--CATCAAT	1111				
Qy	597	AAATAGATTAGTGTATTTGTTCTCTTCAGTACCAATTAACCAATTAAGCAATGCTAGCTT	656				
Db	1117	GGAATAGATTAGTGTATTTGTTCTCTTCAGTACCAATTAACCAATTAAGCAATGCTAGCTT	1051				
Qy	657	ATTGTTTCATATTAGCCATGTTGCACACCCTTAATTGCAACATTA	701				
Db	1057	ATTGTTTCATATTAGCCATGTTGCACACCCTTAATTGCAACATTA	1013				
RESULT 3							
ABZ10246/c							
ID	ABZ10246	standard; DNA; 8056 BP.					
XX							
AC	ABZ10246;						
XX							
DT	16-JAN-2003	(first entry)					
XX							
DE	Haematopoietic cell proliferation disorder related DNA sequence #386.						
XX							
KW	Human; haematopoietic cell proliferation disorder; cytostatic;						
KW	gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;						
KW	cytosine methylation state; gene; ds.						
XX							
OS	Homo sapiens.						
XX							
PN	WO200277272-A2.						
XX							

03-OCT-2002.
XX PF 26-MAR-2002; 2002MO-EP003401.
XX PR 26-MAR-2001; 2001US-0278333P.
PA (EPIG-) EPIDEMIOLOGICS AG.
XX Berlin K, Braun A, Distler J, Gueinig D, Howe A, Mueller J,
PI Olek A, Pispenerbrock C, Adorjan P, Grabs G, Lesche R, Leu E,
PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C,
PI Schwope I, Ziebarth H;
DR WPI; 2003-018942/01.
XX
XX Detecting and differentiating between hematopoietic cell proliferative
PT disorders, comprises contacting a target nucleic acid with a reagent that
PT distinguishes between methylated and non-methylated CpG dinucleotides.
PS Claim 28; SEQ ID NO 386; 117pp; English.
XX
XX The present invention describes a method for detecting and
CC differentiating between haematopoietic cell proliferative disorders
CC associated with at least 1 gene and/or their regulatory regions in a
CC subject. The method comprises contacting a target nucleic acid in a
CC biological sample obtained from the subject with at least 1 reagent,
CC which distinguishes between methylated and non-methylated CpG
CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ1118
CC represent specifically claimed nucleotide sequences from the present
CC invention. Oligonucleotides from the present invention can be used for
CC differentiating between healthy haematopoietic cells and proliferative
CC disorder haematopoietic cells; for differentiating between acute
CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
CC determining the cytosine methylation state and/or single nucleotide
CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
CC related sequences and their complements; and as primers for the
CC amplification of haematopoietic cell proliferation disorder related DNA
CC sequences. The nucleotide sequences from the present invention can also
CC be used for detecting a predisposition to, differentiation between
CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
CC haematopoietic cell proliferative disorders. The present method enables a
CC highly specific classification of haematopoietic cell proliferative
CC disorders allowing for improved and informed treatment of patients
XX
SQ Sequence 8056 BP; 3711 A; 0 C; 371 G; 3974 T; 0 U; 0 Other;

Query Match 4.9%; Score 72; DB 7; Length 8056;
Best Local Similarity 42.0%; Pred. No. 0.00051;
Matches 489; Conservative 0; Mismatches 670; Indels 5; Gaps 1;

OY 203 TTGAACCTTACAGCAGCAATTAATTCGGATTAACCTGCCATAATTAACAATTCTTGAAAC 262
DB 1640 TTTTACAATTCATTTAATTGCATTTTTTAAAAATATTAATAATATATTAATTAATTAAT 1581
OY 263 ATACGATATGTAATCTTTCTTCAATTAAGTGAATTTTGTCT-----TTTTTTAAGCTA 317
DB 1580 ATTAAATTTAAATATATATATTTTTTAAAAAAAATATATTTTAAAAATTTTTTTAACATAT 1521
OY 318 TGACAAATTTGAAAAAAAAGAAATGAAAGTAGAGTGGCTTAACCATTTGAAATATAT 377
DB 1520 TTAATTTTAAATTTTAAATATATATTAATTAATTAATTAATTAATTAATTAATTAATTTAT 1461
OY 378 AGGCTAAGTTTTTCTGATGCGTTTAACATAAAAAAGAAATACAAAAGCTATTAGCAT 437
DB 1460 AAAAAATTAATTTTAAATTTTTTTTTTAAAAAAAATTAATAAAAAATTAATAAAAAATC 1401
OY 438 AACCTCGTAAGTGTCACAAAAATATTTTGGACGTAGCTATAGAAAATATACAA 497
DB 1400 AAATATATTTTAAATAAAAAAAATTAATTTTTTAAATTAATTAATTAATTAATTT 1341
OY 498-ACTTAATCTTAAAGAATTTCTCTATATATATAGGAATCCCTCTCACGTGAAGCA 557
DB 1340 TTATATTTAAATTAATTAATTAATTTTTTAAATTTTAAATTTTAAATTAATTAATTTCAAT 1281

QY	558	ATTATTCATCGAATTATGACGTCACATTCATCCATCAATATAATAGATTTAGTGATTTGTT	6117
Db	1280	TTTATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT	12211
QY	618	CTCTTCAGTACAAATTAATCTACATTAATGCAATGCTAGCTTATTTGTCATTAATACCAATG	677
Db	1220	AAAAAAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	11611
QY	678	TGCAACCCATTAATCGAATCAATTAATGATTCATTAATTTTCTTGCCCTCTTGTTT	737
Db	1160	ATTATATTTTAAATTTTCGAATTTCAAAAAAAATTTAAAAAACAATATATATACATTTTA	11011
QY	738	TCTAACAAAAATGTTCCAGATTTTTPAAAAATTTTGAAAAATACATTAACACTTTGA	797
Db	1100	AAACATTTAAAAAAAACAAATTTAAAAAATAATTTTATTTTACAAATTAATTTTCA	10411
QY	798	GTAATGATTAATTCGAATTAATGACTGTGTTTGAAAGTAAAGATCAAAATTTTCTTACT	857
Db	1040	AAATTTAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	981
QY	858	CGACTTAATGCACTTTACATCAATCGATGATGATTTGATCTATCTATTAAGCTTAACA	917
Db	980	ATTTTATTCATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	921
QY	918	AATTAATCTTTCAAAATGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	977
Db	920	CAACCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	861
QY	978	AAGCAATTCGGAATTTCTGACGATTAATGTCGACTCAATTCATTTGTAATTAAGGAGAG	103
Db	860	ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	801
QY	1038	TTTGTGTGAGCTCATTAATGCTTTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT	109
Db	800	AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	741
QY	1098	CTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	115
Db	740	TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	681
QY	1158	AGCTGTTTTTTTTTGCCTTAATTTTAAGCATCACTTTGTTGACTTTTGTGAGCT	121
Db	680	ATTTTGTTTTTTTTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	621
QY	1218	TTAATACCGTTTTCGACATTTCTCTTGTAATCACTGTAATCGCCTTTTAAACATGACA	127
Db	620	TTAATTAATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	561
QY	1278	ACTCTTGTAAGTCCCTTTCTTTGCCCATTAATTAATCATCTTGGAATATGTAATCAGA	133
Db	560	ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	501
QY	1338	ATAGTTTTTCAAAAATCTAATTAATA 1361	
Db	500	KATTTAATTTTATTTTAAATA 477	
RESULT 4			
ABL32426/c			
ID ABL32426 standard; DNA; 9810 BP.			
XX ABL32426;			
XX AC			
XX DT 26-MAR-2002 (first entry)			
XX DE Human immune system associated gene SEQ ID NO: 399.			
XX KM Human; immune system disease; cytosine methylation; antiaesthetic;			
KM antiarteriosclerotic; antianaemic; cytosolic; noctropic;			
KM neuroprotective; anti-HIV; anticonvulsant; ophthalmological;			
KM antineuritic; antirheumatic; antidiabetic; antipsoriatic;			
KM antineuroinflammatory; cancer; eye disease; arteriosclerosis; anaemia;			